AD-A077 320 DECISIONS AND DESIGNS INC MCLEAN VA F/G 5/10
EVALUATION AND INTEGRATION OF IMPRECISE INFORMATION.(U)
OCT 79 P J STICHA , J J WEISS , M L DONNELL N00014-78-C-0651
UNCLASSIFIED PR-79-21-90 NL

1 OF 1
AD
A077320

END
DATE
FILMED
1-80
DDC

October 1979

FINAL TECHNICAL REPORT PR 79-21-90

12

Evaluation and Integration of Imprecise Information

Paul J. Sticha
Jonathan J. Weiss
Michael L. Donnell

DECISIONS and DESIGNS, INC.

Prepared for
Office of Naval Research
Contract No. N00014-78-C-0651

9 14

FINAL TECHNICAL REPORT PR-79-21-90

1 Aug 78 – 29 Aug 79

6

EVALUATION AND INTEGRATION OF IMPRECISE INFORMATION

by

10

Paul J. Sticha, Jonathan J. Weiss, and Michael L. Donnell

Prepared for

Office of Naval Research
Contract N00014-78-C-0651

15

11

October 1979

12 78

DECISIONS and DESIGNS, INC.

Suite 600, 8400 Westpark Drive
P.O.Box 907
McLean, Virginia 22101
(703) 821-2828

390 664

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER PR 79-21-90	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EVALUATION AND INTEGRATION OF IMPRECISE INFORMATION		5. TYPE OF REPORT & PERIOD COVERED Final Report 1 Aug 1978-29 Aug 1979
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Paul J. Sticha Jonathan J. Weiss Michael L. Donnell		8. CONTRACT OR GRANT NUMBER(s) Contract No. N00014-78-C-0651
9. PERFORMING ORGANIZATION NAME AND ADDRESS Decisions and Designs, Incorporated Suite 600, 8400 Westpark Drive, P.O. Box 907 McLean, VA 22101		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, VA 22217		12. REPORT DATE October 1979
		13. NUMBER OF PAGES 79
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Fuzzy Set Theory
Psychological Measurement & Scaling
Reasoning
Logic
Mathematical Psychology
Psychophysics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Although fuzzy set theory has been rapidly gaining popularity as a rigorous framework for incorporating imprecision into quantitative reasoning, a brief survey of the existing literature indicates the lack of any consistently applied operational definition for the fundamental concept of membership function. In the present case, the membership of an element x in a set S is defined as the degree of truth of the statement "x is a member of S." It then becomes (con't)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont

necessary to develop an empirically valid scale of truth which allows not only the binary extremes of "true" and "false," but also the continuum of intermediate values. In one experiment,

In the first of two experiments, subjects performed two tasks: pair-wise comparison; and direct numerical scaling of the relative truth of simple sentences. Results indicated that (1) the high degree of transitivity in each subject's paired-comparison judgments leads us to reject the hypothesis of a two-valued true-false logic in favor of a continuum of values; (2) ability to discriminate, as judged by the consistency between direct ratings and paired-comparison judgments, seems to be uniform along the true-false continuum, again favoring the hypothesis of a continuum of truth values over that of a binary categorical judgment; and (3) the high correlation between an item's aggregate binary preference score for a given subject and that subject's direct rating for the item indicates that at least two different methods of inferring degree of truth are highly consistent.

In the second experiment, subjects estimated the truth of conjunctions and disjunctions of simple sentences using methods of rank order and direct estimation. Analysis of both ordinal data and direct estimates indicates that the truth of conjunctions or disjunctions of sentences is a multiplicative function of the truth of the constituent sentences, allowing these functions to be used to obtain cardinal scale information about subjective truth. In addition, the possibility of response biases in direct estimation of the truth of sentences of different degrees of complexity was identified. Both descriptive and normative implications of these results were discussed, as well as implications for future research.

CONTENTS

	Page
DD FORM 1473	ii
FIGURES	vi
TABLES	vii
1.0 INTRODUCTION	1
2.0 REVIEW OF EMPIRICAL LITERATURE	3
2.1 Methods of Assessment	3
2.2 Experimental Findings	5
3.0 EXPERIMENT 1	8
3.1 Method	9
3.1.1 Subjects	9
3.1.2 Stimuli	9
3.1.3 Procedure	10
3.2 Results	11
3.2.1 Transitivity of paired-comparison judgments	11
3.2.2 Categorical perception of truth	17
3.2.3 Consistency of paired-comparison choices and direct estimates	18
3.3 Discussion	21
4.0 EXPERIMENT 2	23
4.1 Introduction	23
4.1.1 Ordinal conditions for Min- and Max-representations	25
4.1.2 Consistency of Min- and Max-representations	33
4.1.3 Ordinal conditions for product and inverse product representations	34
4.1.4 Consistency of product and inverse product representations	36

CONTENTS (Cont'd)

	Page
4.2 Method	37
4.2.1 Subjects	37
4.2.2 Stimuli	37
4.2.3 Procedure	39
4.3 Results	41
4.3.1 Consistency	41
4.3.2 Analysis of ordinal data	41
4.3.3 Analysis of direct estimates	45
4.4 Discussion	51
5.0 DISCUSSION AND CONCLUSIONS	56
REFERENCES	59
APPENDIX A. LIST OF STIMULUS SENTENCES	A-1
APPENDIX B. A BRIEF EXPLANATION OF "STEM-AND-LEAF" DIAGRAMS	B-1
APPENDIX C. DISTRIBUTION LIST	C-1

FIGURES

Figure		Page
3-1	Test of Categorical Perception of Truth	19
4-1	Indifference Curves for F(x,y) = Min(x,y)	27
4-2	Lattice Points in Single Factorial Experimental Design	28
4-3	Illustration of Scaling Procedure for Min Representation	32
4-4	Illustration of Compatibility of Min and Max Representations	35

TABLES

Table		Page
3-1	Stem-and-Leaf Display of Number of Intransitive Triples	12
3-2	Consistency Measures for Paired-Comparison Data	13
3-3	Stem-and-Leaf Display of Inferred Probability of Pair Reversal	16
3-4	Stem-and-Leaf Display of Correlation Between Paired-Comparison Score and Direct Truth Estimate	20
4-1	Sentences Used to Construct Experimental Stimuli	38
4-2	Consistency of Rank Order and Direct Estimates	42
4-3	Pairwise Reversals in Rank Order Necessary to Satisfy Representation	44
4-4	Mean Squared Deviations From Model Predictions Taken from Simple Estimates	47
4-5	Summary of Analyses of Variance for Direct Estimates of Conjunctions	49
4-6	Summary of Analyses of Variance for Direct Estimates of Disjunctions	50

EVALUATION AND INTEGRATION OF IMPRECISE INFORMATION

1.0 INTRODUCTION

This report summarizes the empirical research conducted by Decisions and Designs, Incorporated (DDI) on the psychological processes by which individuals evaluate and integrate imprecise information. Organized around the theory of fuzzy sets, the research reported here attempts to determine how well judgments about partial truth of simple or compound statements correspond with the numerical structure of fuzzy set theory. This research involves determining empirical conditions implied by formulations of fuzzy set theory, and testing whether human judgments satisfy these conditions; its ultimate goal is to increase understanding of the human ability to interpret imprecise information, and to develop insight into the measurement and scaling of concepts related to fuzzy set theory, such as set membership and partial truth.

Fuzzy set theory is a comparatively new field, originally developed by systems engineers as a method of dealing approximately with highly complex systems. The progress made during the initial decade of fuzzy set development has been strongest, therefore, in the areas of practical implementation and mathematical representation. However, in the area of empirical behavioral validation of the theory, the literature has been sparse, incomplete, and inconsistent. Because the mathematical properties of a system that would satisfy our naive intuitive definitions are not uniquely determined, and because the goals and backgrounds of fuzzy set researchers differ widely, there have been to date no definitive, standard set of notations and formulas, and no

consistent framework in which to compare and reconcile results from different researchers. This inconsistency has been somewhat of an embarrassment to those attempting to assimilate the new technology, often leading to the hasty conclusion that the concept of fuzziness is inherently inconsistent or illogical. The present studies are designed to demonstrate that an operational, empirically valid set of definitions, formulas, etc., which can act as a unifying framework for fuzzy set theory, does exist.

The following sections of this report describe the research conducted by DDI in detail, and its theoretical and practical implications. Section 2.0 contains a brief survey of the current literature concerning empirical psychological studies of reasoning with imprecisely defined concepts. Section 3.0 describes the first experiment, which investigates the transitivity of paired comparison judgments concerning the relative truth of simple sentences. Section 4.0 describes the second experiment, in which the structure of judgments about the truth of more complex sentences is investigated. Finally, Section 5.0 provides a general discussion of what the results of this research imply about the state of fuzzy set theory and the conclusions that can be drawn about future research in the area.

2.0 REVIEW OF EMPIRICAL LITERATURE

Despite much recent theoretical interest in fuzzy set theory and frequent assertions that human reasoning may be organized by imprecise rules, there has been little empirical research to verify the consistent use by individuals of fuzzy reasoning processes. The little research already conducted does not present a coherent picture of the human as a processor of fuzzy logical information because there has been no agreement about the types of reasoning tasks to which fuzzy set theory applies, and because basic questions about consistency in the use of imprecise information have not been addressed. This section reviews the limited empirical literature testing fuzzy set theories of human reasoning, paying particular attention to the points mentioned above.

2.1 Methods of Assessment

Although the axioms of fuzzy set theory have the potential for application to a variety of human activities, care must be taken that results in one area are not generalized unjustifiably to different reasoning tasks. Recent empirical research has investigated fuzzy reasoning in a variety of contexts. Graded-set membership functions have been inferred from confidence judgments (Macvicar-Whelan, 1978), choice proportions over individuals (Hersh & Caramazza, 1976), paired comparison and direct judgments of degree of truth of statements (Oden, 1977a, 1977b), degree of agreement with statements (Dreyfuss, Kochen, Badre, & Robinson, 1975), judgments of how well objects exemplify concepts (Rosch, 1975), and estimates of the range of possibility (Hersh, Note 1), the fuzzy analogue to the statistical concept of a confidence interval.

Potentially, of course, any task provides a set of behavioral data which may be modeled by fuzzy set theory. However, not all tasks pertain to questions of whether individuals perceive environmental partial truth or have imprecise internal representations of truth or set membership. For example, Macvicar-Whelan (1978) inferred membership functions from confidence judgments on statements about height. Although inferences were made about the representation of height, the judgments had nothing at all to do with height; rather, they were measures of confidence. Individuals may be more or less confident about statements, even if all their perceptions and internal representations are precise. Fuzzy set theory may provide a model of confidence, but this experiment does not tell us how we understand sentences about height.

A similar problem occurs when membership functions are inferred from group response proportions. This method was used by Hersh and Caramazza (1976), who had subjects judge the applicability of statements about the sizes of various squares. The membership function was taken to be the proportion of "yes" responses within the group. Rationale for this definition is Zadeh's (1968) assertion that the probability of an event is the expected value of its membership function. However, this assertion cannot be used to infer individual membership functions from group response probabilities. The probability merely provides a measure of the degree of consensus among respondents. Furthermore, additional assumptions which allow the inference of individual membership functions must deny individual differences in membership functions. A somewhat similar method was used by Thole, Zimmermann, and Zysno (1979), who inferred differences in set membership from group choice proportions.

Other investigations have been somewhat more direct in their measurement of individual membership functions.

However, since there seems to be a tendency for investigators to use personal measurement schemes, it is important to avoid overgeneralizing results of a particular scheme to all schemes. The reader is advised to pay careful attention to measurement techniques when considering the results presented below.

2.2 Experimental Findings

Researchers seeking some empirical verification for fuzzy set theory have been principally concerned with three questions:

1. Is it possible to obtain quantifiably imprecise judgments from individuals, or other judgments from which the degree of imprecision may be inferred?

2. Does individual understanding of the logical operations ("not," "and," "or," etc.) correspond to the conjectures of Zadeh (1965) and others?

3. What function is served by linguistic hedges, such as the word "very"?

It should not be at all surprising that individuals occasionally produce imprecise judgments. What is surprising is that this fact has been used as evidence for the "psychological reality of fuzzy sets." Macvicar-Whelan (1978), using confidence judgments, found imprecise boundaries between those squares judged "large," for example, and those not judged "large." Kochen and Badre (1974), and Dreyfuss et al. (1975) investigated the effect of context on imprecision in judgment, and found that context information tends to increase judgmental precision when context information is related in a sensible way to the judgment, while nonsensical context information decreases precision.

Zadeh (1965) has suggested the following representation for the logical functions negation, conjunction, and disjunction: The truth value of the negation of a statement is assumed to be the complement of the truth value of the statement; that is, the sum of the truth values of a statement and its negation is 1. The truth value of the conjunction of two statements is the minimum of the truth values of the individual statements. The truth value of the disjunction of two statements is the maximum of the truth values of the two statements. Alternative formulations have been suggested by others (Goguen, 1969), and no consensus has yet been reached as to which of these formulations is "correct."

Hersh and Caramazza (1976) have investigated negation by using group probabilities and confidence judgments to assess truth functions. The data they obtained seem to support Zadeh's conjecture. However, the results should be interpreted in light of the assessment method used. Oden (1977b) used functional measurement techniques (Anderson, 1974) to investigate the role of conjunction and disjunction in imprecise reasoning. The results indicate that conjunction may be represented by the product of individual truth functions, and disjunction by the sum of the truth values less their product. Judgments of truth for conjunctions and disjunctions directly contradicted the minimum and maximum functions suggested by Zadeh. On the other hand, Thole et al. (1979) found evidence supporting the minimum operation for set intersection.

The linguistic hedge "very" may have two functions: It may serve to translate the boundary between true and false sentences; or, it may serve as an intensifier, making judgments more precise. Kochen and Badre (1974) found evidence for the use of "very" as an intensifier. Macvicar-Whelan (1978), and Hersh and Caramazza (1976), on the other hand, found that "very" did not serve to increase precision.

The empirical literature presents an incomplete picture of individuals' usage of imprecise reasoning. In particular, there is some confusion about the types of tasks in which imprecise heuristics are used. There are also some conflicting results on certain aspects of imprecise reasoning. Most important, basic questions remain unanswered, the most fundamental of which is consistency. That is, do people respond to imprecision in a consistent and coherent manner? And by what rules are complex judgments derived from their simpler constituents? The question of consistency, therefore, is the principal concern of DDI's current investigation.

3.0 EXPERIMENT 1

The first experiment in DDI's research effort was designed to investigate the consistency with which individuals make judgments about the truth of sentences. Only individual results were used, because it is important to distinguish the true fuzziness inherent in a subject's knowledge and beliefs from the less interesting phenomenon of "noise" due to the aggregation of subjects whose perceptions do not coincide. Two methods of measurement were used: Subjects first rated the relative truth of sentences in paired-comparison judgments; later, they made direct numerical estimates of the degree of truth attributed to each sentence. Consistency was tested with respect both to the transitivity of the paired comparisons, and to the compatability of those paired judgments with the direct scaling data.

In interpreting the results of this and subsequent experiments, the reader should be aware of the natural correspondence between truth value and set membership. In standard two-valued logic, a statement is either "true" or "false," just as in classical set theory an element is either a "member" or a "nonmember" of a set. Thus, the truth value of the statement, "$\underline{x}$ is an element of the set $\underline{S}$," is by definition equivalent to the <u>degree</u> <u>of</u> <u>membership</u> of $\underline{x}$ in $\underline{S}$. Similarly, the truth value of any sentence $\underline{Y}$ can be identified with the degree of membership of $\underline{Y}$ in the set of true sentences. When extending these classical two-valued notions to the realm of fuzzy set membership and multi-valued logic, we may simply replace the binary set $\{0,1\}$ by the interval $[0,1]$. Thus, 0 represents the truth value of a "false" statement, or the degree of membership for an item which is completely excluded from a given set; 1 represents the truth value of a completely "true" sentence, or the degree of membership of an item which is unambiguously

a member of a given set; and intermediate values represent sentences which are judged only partly true, or elements which are only partially included in a given set.

Just as "truth" (in the classical sense) and "probability" (in Bayesian statistics) may be viewed as subjective, primitive concepts, we may stipulate that "degree of truth" or "level of membership" is also such a concept. In other words, although there is no a priori way of objectively defining the "real" truth value of a sentence (other than self-contradictions or tautologies), we may productively use the concept of subjectively defined partial truth if we can verify that people deal with it (at least ideally) in the same way. In other words, we assume, in a manner entirely analogous to the classical logician or the Bayesian, that individual differences in assigned truth values represent differences in information states and in personal values, but not fundamental variations in the subjects' essential concept of "degree of truth."

3.1 Method

3.1.1 Subjects - Subjects were 11 male and 14 female employees of DDI. Subjects were not familiar with the purpose of the experiment. The experiment took up to an hour and was conducted during office hours. The subjects received no pay other than their regular salary for their participation in the experiment.

3.1.2 Stimuli - Stimuli were a set of 50 sentences. Sentences were limited to those for which truth could be evaluated, and selected to cover, as uniformly as possible, the range from "true" to "false." Sentences were non-controversial and were written to avoid dependence on subjects' personal values or on special knowledge. Stimuli for any single subject were 20 sentences randomly chosen from

the entire set. Each subject received a different set of sentences. A list of the 50 stimulus sentences appears in Appendix A.

3.1.3 <u>Procedure</u> - The experiment had two parts. In the first part of the experiment, subjects viewed all pairs of the 20 sentences, and indicated for each pair the sentence they felt was truer. Subjects were presented with all 190 pairs for a single replication, along with the following instructions:

> Often in real life we encounter statements which seem neither entirely true nor entirely false. Given two of these statements, we may be able to say that one or the other is more true, although we could not say that either is absolutely true or absolutely false.
>
> On the following pages are several pairs of statements labeled A and B. For each pair, circle the letter of the statement you feel is more true.
>
> There are no right or wrong answers to these questions, and this test in no way measures your intelligence or personality. In fact, your responses will not be compared to the individual responses of any other subjects. There are a lot of questions; work quickly but carefully. Your help in this research is much appreciated.
>
> Remember, for each pair, circle the letter of the sentence you feel is more true. Please answer every question, even though it may be difficult on some problems for you to decide which letter to circle. When you finish all sentence pairs, continue with the brief second part of the experiment. If you have no questions, you may begin.

In the second part of the experiment, subjects individually estimated the truth of each sentence, assigning a number between 0 and 100 to each sentence in proportion to its perceived truth. Subjects received the following instructions:

> For this part of the experiment, you will see the same sentences you saw before, now one at a time instead of in pairs.

> Your task is to assign to each sentence a number between 0 and 100 in proportion to how true you think each sentence is. If you think the sentence is completely true, assign the number 100; if you think the sentence is completely false, assign the number 0; otherwise assign an intermediate number corresponding to the sentence's degree of truth. Put the number on the line to the left of each sentence. If you have no questions, you may begin.

Data consisted of paired-comparison judgments and direct estimates for each subject.

3.2 Results

3.2.1 Transitivity of paired-comparison judgments - For each subject's paired-comparison data, the number of intransitive triples was calculated. The distribution of the number of intransitive triples is presented in the stem-and-leaf display (Tukey, 1977) in Table 3-1. (Readers who are unfamiliar with stem-and-leaf displays can find a brief explanation in Appendix B.) The median number of intransitive triples is 17, and the first and third quartiles are given by 8 and 24, respectively. To help interpret these numbers, the maximum possible number of intransitive triples is 330, and the expected number, given random choice on each pair, would be 285. Clearly, there is quite a bit of regularity in the paired-comparison judgments.

Kendall (1955) has shown that under the assumption of random choice, the number of intransitive triples is distributed approximately χ^2, and has derived formulas for calculating a test statistic and appropriate degrees of freedom. The third column of Table 3-2 shows the value of χ^2 for each subject. Examination of this table shows that every subject had significantly fewer intransitive triples than would be predicted by chance, $p < .001$.

Stem-and-Leaf Display of
Number of Intransitive Triples

0	2	3	4	4	5	5	8	8
10	1	3	5	5	7	7	8	9
20	3	4	4	5	8			
30	0	7	9					
.								
.								
.								
.								
110	7							

Table 3-1

Consistency Measures for Paired-Comparison Data

Subject	Number of Intransitive Triples	χ^2 (Randomness) df = 27	χ^2 (True-False) df = 40	χ^2 (Three Group) df = 75
1	6	166.5^{t}	113.3^{t}	127.0^{t}
2	15	162.0^{t}	101.3^{t}	100.0^{f}
3	5	167.0^{t}	114.7^{t}	130.0^{t}
4	39	150.0^{t}	69.3^{h}	28.0
5	18	160.5^{t}	97.3^{t}	91.0
6	28	155.5^{t}	84.0^{t}	61.0
7	23	158.0^{t}	90.7^{t}	76.0
8	6	166.5^{t}	113.3^{t}	127.0^{t}
9	15	162.0^{t}	101.3^{t}	100.0^{f}
10	8	165.5^{t}	110.7^{t}	121.0^{t}
11	24	157.5^{t}	89.3^{t}	73.0
12	2	168.5^{t}	118.7^{t}	139.0^{t}
13	17	161.0^{t}	98.7^{t}	94.0
14	24	157.5^{t}	89.3^{t}	73.0
15	17	161.0^{t}	98.7^{t}	94.0
16	30	154.5^{t}	81.3^{t}	55.0
17	37	151.0^{t}	72.0^{h}	34.0
18	3	168.0^{t}	117.3^{t}	136.0^{t}
19	11	164.0^{t}	106.7^{t}	112.0^{h}
20	19	160.0^{t}	96.0^{t}	88.0
21	8	165.5^{t}	110.7^{t}	121.0^{t}
22	117	111.0^{t}	−34.7	−206.0
23	25	157.0^{t}	88.0^{t}	70.0
24	5	167.0^{t}	114.7^{t}	130.0^{t}
25	13	163.0^{t}	104.0^{t}	106.0^{f}

$^{f}p < .05$ $^{h}p < .01$ $^{t}p < .001$

Table 3-2

A somewhat more interesting null hypothesis than that of random choice is the hypothesis that the set of sentences may be partitioned into two distinct subsets: "true" sentences and "false" sentences. Under this hypothesis, subjects will always say that a "true" statement is truer than a "false" statement, but will choose randomly between pairs of "true" statements or pairs of "false" statements. The expected number of intransitive triples under this hypothesis is smaller than under the hypothesis of random choice. Furthermore, the expected number depends on the number of sentences in the true and false sets. The expected number is lowest when half of the sentences are true and half are false, so this case gives, in some sense, the most stringent test of consistency compared to the true-false hypothesis. In the above case, it is possible to modify the test statistic to test the true-false hypotheses. Values of χ^2 are presented in the fourth column of Table 3-2. The expected and maximum number of intransitivities under the true-false hypothesis are 60 and 80, respectively. As illustrated in Table 3-2, all subjects but one had significantly fewer violations than would be expected under the true-false hypothesis, $p < .001$. The remaining subject had more violations of intransitivity than would be possible under the true-false hypothesis.

It may be plausible to hypothesize that sentences are divided into three groups: "true" sentences, "false" sentences, and those sentences which are neither true nor false or for some reason cannot be evaluated as true or false. It is possible to derive a statistic to test this three-group hypothesis, but its approximation by a χ^2 distribution is only fair. Nevertheless, the fifth column of Table 3-2 shows that 12 of the 25 subjects had significantly fewer intransitive triples than would be predicted by the three-group hypothesis, $p < .05$. These independent χ^2 variables may be added to obtain an overall indication

of consistency. The obtained overall χ^2 of 2,080 is highly significant, $\underline{df} = 1,860$, $\underline{p} < .001$.

To summarize, the consistency of paired comparison responses was much greater than would be expected by either random choice or a two-group hypothesis. For 12 of the subjects, the three-group hypothesis could also be rejected; furthermore, the three-group hypothesis could also be rejected when individual data were aggregated over all subjects.

Another way of assessing the level of consistency is by estimating what value for the probability of making an error on a pair would produce the obtained rate of inconsistency. In other words, suppose subjects were transitive, except that they reversed their choice randomly with probability $\underline{p}$. Then, the probability of an intransitive triple, $\underline{p}_i$ is given by

$$\underline{p}_i = \underline{p}(1 - \underline{p}) \quad (1)$$

If the obtained proportion of intransitive triples is $\underline{p}_o$, then $\underline{p}$ may be estimated by

$$\hat{\underline{p}} = \frac{1 - \sqrt{1 - 4\underline{p}_o}}{2} \quad (2)$$

Note that the expected proportion of intransitive triples reaches its maximum when $\underline{p}$ is .5.

The obtained distribution of $\hat{\underline{p}}$ is displayed in the stem-and-leaf display in Table 3-3. This distribution ranges from .18% to 11.61% and has a median of 1.51%. For all subjects except one, the inferred probability of paired comparison reversal is less then 3.6%. Clearly, the

Stem-and-Leaf Display of

Inferred Probability of Pair Reversal

.00	2	3	4	4	5	5	7	7
.01	0	2	3	3	5	5	6	7
.02	1	2	2	2	5	7		
.03	4	6						
.04								
.								
.								
.								
.11	6							

Table 3-3

deviations from transitivity are minor, and consistent with a reasonable level of experimental error.

3.2.2 Categorical perception of truth - It seems clear that the subjects consistently respond to the truth differences between sentences. However, the sentences may still be perceived to be either true or false. If this is the case, then subjects may have more difficulty in distinguishing the relative truth of two sentences both perceived as true or both perceived as false, than in distinguishing the relative truth of a true and a false statement. This type of result has been used as evidence for categorical perception of speech sounds (Liberman, Harris, Hoffman, & Griffith, 1957); sounds which are placed in different categories are much easier to distinguish than sounds placed in the same category, even if the sounds in different categories are much more similar on some physical dimension than the sounds in the same category.

A fundamental difference between the sentences used in this investigation and computer-generated speech sounds is that there is no environmental measure of partial truth. However, it may seem reasonable to use the direct truth ratings as the objective measure of truth. Since the direct ratings were on a scale of 0 to 100, statements rated greater than 50 were interpreted as "true" statements; those rated less than 50 were interpreted as "false" statements. The statement pairs were then divided into four groups. The "true" group included those pairs in which both statements were "true," or one was "true" and the other was rated 50. The "false" group included those pairs in which both statements were "false," or one was "false" and one was rated 50. The mixed group included those pairs consisting of a "true" and a "false" sentence. The remaining pairs, consisting of two sentences rated 50, were not analyzed further (there were 65 such pairs, or 1.37% of the total sample).

For each group, the probability of choosing the sentence with the higher directly estimated truth value was calculated as a function of the difference in the estimated truth values. Direct estimate differences were grouped by 10's so that groups consisted of pairs with differences between 1 to 10, 11 to 20, and so forth. If the perception of truth is categorical, the proportion of choices consistent with direct estimates should be greater for the mixed group than for the true or false group for any scale difference.

The results of this analysis are displayed in Figure 3-1. The obtained proportions are displayed along with 95% confidence intervals for the actual probability. Examination of Figure 3-1 shows that the probability of a choice consistent with the direct estimates indeed increases when the difference in the estimates increases. There appear to be no differences between the mixed group and the other two groups. An analysis of variance of the transformed proportions for the range of scale differences from 1 to 50 confirms what seems obvious from Figure 3-1. The effect of difference in scale value was large and significant, $\underline{F}(4,8) = 9.99$, $\underline{p} < .01$, accounting for 81% of the total sum of squares. The effect of the pair group was not significant, $\underline{F}(2,8) = 0.67$, accounting for only 2.7% of the total sum of squares. Thus, there is no evidence for categorical perception of truth.

3.2.3 <u>Consistency of paired-comparison choices and direct estimates</u> - The results described in the previous section suggest that there is substantial agreement between the two methods of assessing partial truth. To check this directly, the product-moment correlation was calculated for each subject between the number of statements a sentence was judged truer than in the paired judgments and the sentence's directly estimated truth value. A stem-and-leaf display of the distribution of correlations is shown in Table 3-4. The

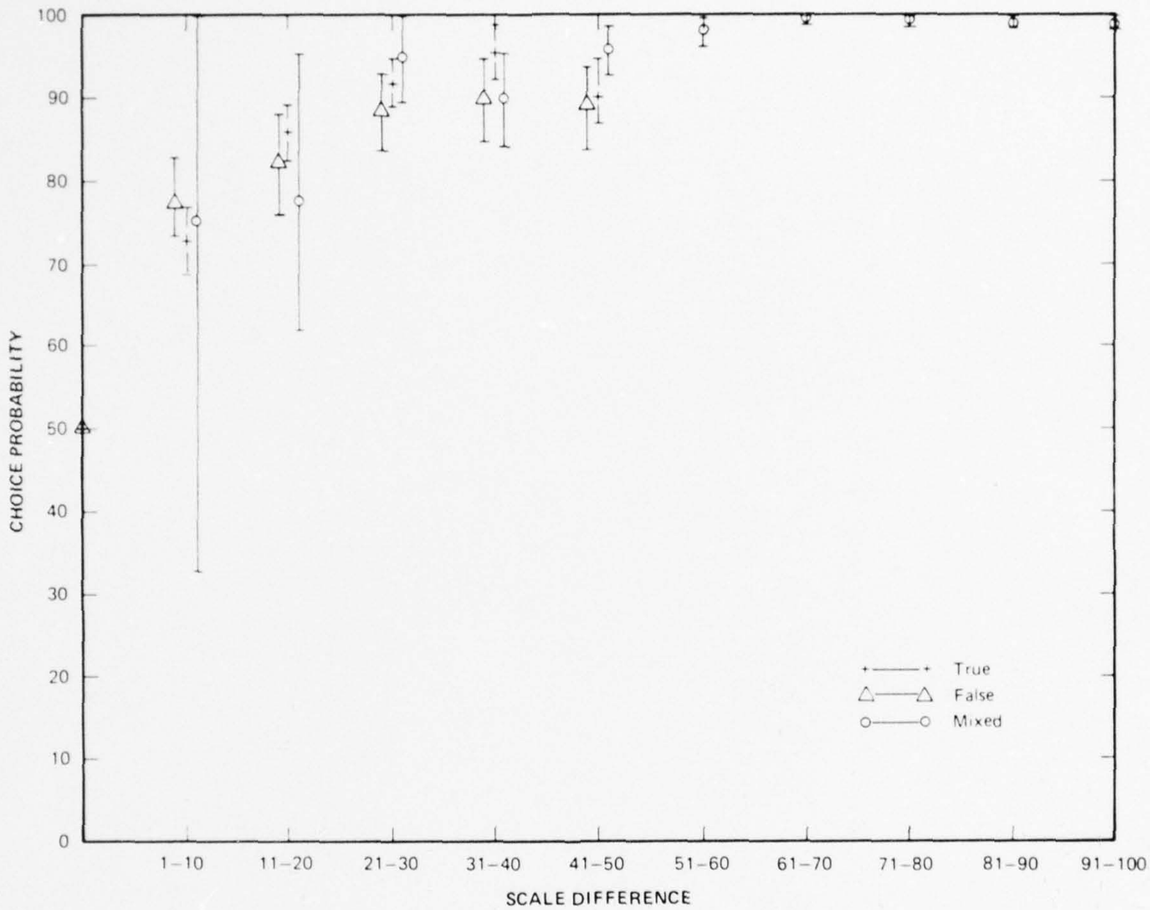


Figure 3-1

TEST OF CATEGORICAL PERCEPTION OF TRUTH

Stem-and Leaf Display of Correlation Between
Paired-Comparison Score and Direct Truth Estimate

.700 – .749	37, 41
.750 – .799	76
.800 – .849	20, 46
.850 – .899	55, 58, 87, 88, 93, 94, 97
.900 – .949	00, 01, 02, 10, 12, 16, 19, 37, 38
.950 – .999	54, 59, 70, 77

Table 3-4

correlations ranged from .737 to .977, with a median of .900. Clearly, there is almost complete agreement between the two assessment methods.

3.3 Discussion

In order to represent judgments of relative truth by an ordinal scale, it is necessary that these judgments be connected and transitive. The data seem unequivocal on transitivity. There were far fewer violations of transitivity than would be expected if subjects responded to sentences as though they were either true or false. The results also seem to indicate that individuals perceived the truth of the statements in a continuous rather than a categorical manner. Furthermore, two measurement techniques gave highly similar truth scales. Although the results seem clear, attention should be paid to limitations in the experiment.

There are two differences between the present analysis of categorical perception and those which have been done in other areas (e.g., Liberman et al., 1957). First, in the current analysis, there is no objective measure of partial truth. Since two subjective measures are being compared, both may be perceived in the same way, and no categorization effect will be found regardless of the nature of subjects' perceptions. Second, categorical perception has been tested in a single context only. In the perception of speech sounds, the main strength of the findings that perception is categorical comes from the fact that in a different context, when individuals employ a physical rather than a linguistic interpretation of the sounds, perception is continuous. The existence of contextual dependencies corrects for possible confounding effects of nonlinear subjective scales. This result suggests that although perception of truth may be continuous in the context in which it was analyzed, it may be that categorical perception would occur in other contexts,

such as in more complex judgment and decisionmaking tasks. Further research in this direction may serve to illuminate this problem.

It must be realized that both the task used and the type of consistency tested are somewhat limiting. Behavior in a simple paired-comparison task will not necessarily reflect the same processes as the behavior in a more complex decisionmaking task. Furthermore, the statements themselves were simple. A much stronger test of consistency would involve sentences formed by joining two or more simple statements with appropriate logical connectives. Tests of this stronger form of consistency would clarify the processes by which the truth of compound statements is evaluated, give a stronger representation of partial truth, and further verify the validity of partial truth as a factor in human reasoning.

4.0 EXPERIMENT 2

4.1 Introduction

The results of Experiment 1 indicate that it is possible to obtain an ordinal representation of partial truth. This second experiment seeks to strengthen this representation, and to test those conditions which would lead to an ordered metric or interval scale of partial truth. In order to accomplish this goal, the truth of compound sentences, joined by logical operations of conjunction or disjunction, are assessed--both ordinally and by direct estimation.

Cardinal measurement scales for properties of objects are typically obtained by the use of some rule for combining objects. Examples of such rules include the concatenation operation in extensive measurement; difference or similarity judgments; probability combinations of outcomes in utility theory; and the ordering of cells in the cross product of two sets in conjoint measurement. The cardinal scale is obtained by assuming that the combination operation satisfies some kind of cancellation axiom. For example, in extensive measurement of mass, the cancellation axiom states that adding the same mass to each of two existing masses does not change which of the two is heavier. This cancellation axiom is the principal condition to be satisfied in order to obtain a cardinal measure of mass. Other conditions are also necessary, however. In particular, a technical condition called solvability asserts that the set of objects is sufficiently dense to guarantee that certain equations can be solved. Although this condition cannot be tested empirically, it is critical for obtaining an interval or higher scale, rather than an ordered metric scale. Thus, the same empirical conditions underly both ordered metric scales and cardinal scales; the major difference between the levels of uniqueness is in the density of the set of objects.

The most natural operations for combining fuzzy sets are the operations of union and intersection, and the corresponding logical operations of disjunction and conjunction. There are several reasons why these operations provide a good basis for investigation in imprecise reasoning. First, the operations are basic to set theory; all set theoretic operations may be formed by combinations of these operations and complementation, so that representation of fuzzy set membership based on these operations would completely characterize other, more complicated, operations. Second, there is some current controversy and conflicting empirical evidence about the form of the function which should represent these operations. As mentioned earlier, Zadeh (1965) has suggested that intersection and union be represented by the functions Min and Max, respectively; alternatively, these operations may be represented by the product and inverted product, respectively. Third, the form of the operations has implications for the measurement of partial truth. In particular, if Min and Max accurately represent judgments of conjunction and disjunction, then an ordinal representation of truth is sufficient for all applications. (To be more precise, the scale is somewhat more restricted than an ordinal scale because it must satisfy two additional properties: The minimum and maximum values of the scale must be 0 and 1, respectively; and, the sum of the truth of a statement and its negation must be 1. However, any monotonic function may be transformed to satisfy the above conditions.) If, however, the product and inverse product, or some other compensatory rules represent the truth of conjunctions and disjunctions, then judgments about the relative truth of compound sentences may be used to obtain ordered metric or interval information about the truth scales.

The task of this experiment, therefore, is twofold: The primary task is to test alternative representations of

the subjective truth of conjunctions and disjunctions; the second task is to determine the measurement and scaling implications of these representations, and to find a scale of truth consistent with the representations. To accomplish this task, subjects estimated the truth of pairs of sentences joined by "and" or "or." Examples of these sentences are the conjunction, "Four inches [10.16 cm] is a heavy snowfall, and Chevrolet makes excellent cars," and the disjunction, "Either four inches [10.16 cm] is a heavy snowfall, or Chevrolet makes excellent cars, or both." Two methods of estimation were used. In the first method, subjects ranked the 16 possible combinations formed by combining a sentence from each of two sets of four sentences according to their relative truth. In the second method, subjects directly estimated the truth of the same sentences by assigning to each a number between 0 and 100.

The general analysis strategy for the data is to assess the extent to which the ranking or direct estimates may be predicted by the two appropriate numerical functions. For the direct estimates, the assessment of fit is accomplished by more or less standard techniques in which deviations from best-fitting predictions are examined. For the ordinal data, the analysis involves the application of techniques such as additive conjoint measurement. There are two questions, however, which have not been addressed in conjoint measurement that are important in this investigation. The first is the conditions which are necessary and sufficient for a rank ordering to be predicted by a Min or Max function. The second question is the consistency of two rank orders with the same truth scale.

4.1.1 <u>Ordinal conditions for Min- and Max-representations</u> - This section identifies the conditions under which a rank order on the cross product of two finite sets may be

described by a minimum or maximum rule. More precisely, we define the following terms.

Definition. Let $\underline{A}_1 = \{\underline{a}_{11}, \underline{a}_{12}, \ldots, \underline{a}_{1\underline{m}}\}$ and $\underline{A}_2 = \{\underline{a}_{21}, \underline{a}_{22}, \ldots, \underline{a}_{2\underline{n}}\}$, and let $\succsim$ represent a weak ordering on $\underline{X} = \underline{A}_1 \times \underline{A}_2$. $\succsim$ is Min-representable if there exist $\underline{k}_1, \underline{k}_2, \ldots, \underline{k}_{\underline{m}}$ and $\underline{\ell}_1, \underline{\ell}_2, \ldots, \underline{\ell}_{\underline{n}} \in \mathrm{Re}$ such that $\underline{x}_{\underline{ij}} \succsim \underline{x}_{\underline{i}'\underline{j}'}$ whenever $\mathrm{Min}(\underline{k}_{\underline{i}}, \underline{\ell}_{\underline{j}}) \geq \mathrm{Min}(\underline{k}_{\underline{i}'}, \underline{\ell}_{\underline{j}'})$, where $\underline{x}_{\underline{ij}} = \langle \underline{a}_{1\underline{i}}, \underline{a}_{2\underline{j}} \rangle$.

It should be noted that when the minima are equal, either one ordering or its reverse is consistent with the representation.

It is possible to define Max-representable in a completely analogous fashion, merely by substituting Max for Min in the definition.

It is now necessary to specify a set of empirical conditions which guarantee that a rank order is Min-representable (the conditions may be modified appropriately to obtain a set of empirical conditions sufficient for an ordering to be Max-representable). The motivation for these conditions may be clarified by an examination of the indifference curves for the function $\underline{F}(\underline{x},\underline{y}) = \mathrm{Min}(\underline{x},\underline{y})$, as illustrated in Figure 4-1. These curves are horizontal lines in the region below the line $\underline{y} = \underline{x}$, and vertical lines in the region above the line $\underline{y} = \underline{x}$.

The points represented by the cross product of two finite sets may be represented by a lattice of points as in Figure 4-2. In this figure, all points on the bottom row lie on a single indifference curve; these points have a lower minimum than any other points in the lattice. Within that row, all points are equivalent; any arbitrary ordering

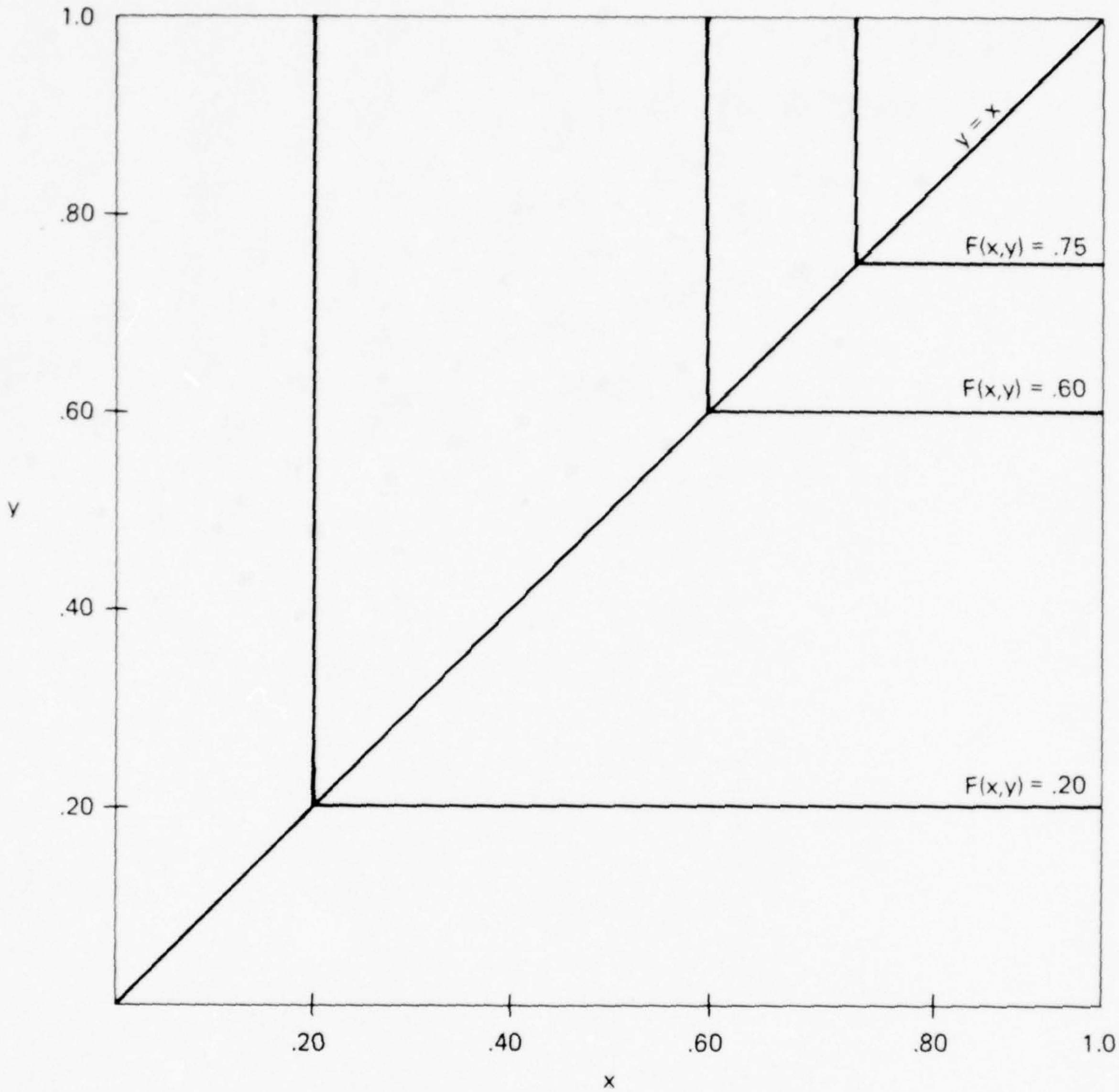


Figure 4-1
INDIFFERENCE CURVES FOR $F(x,y) = \min(x,y)$

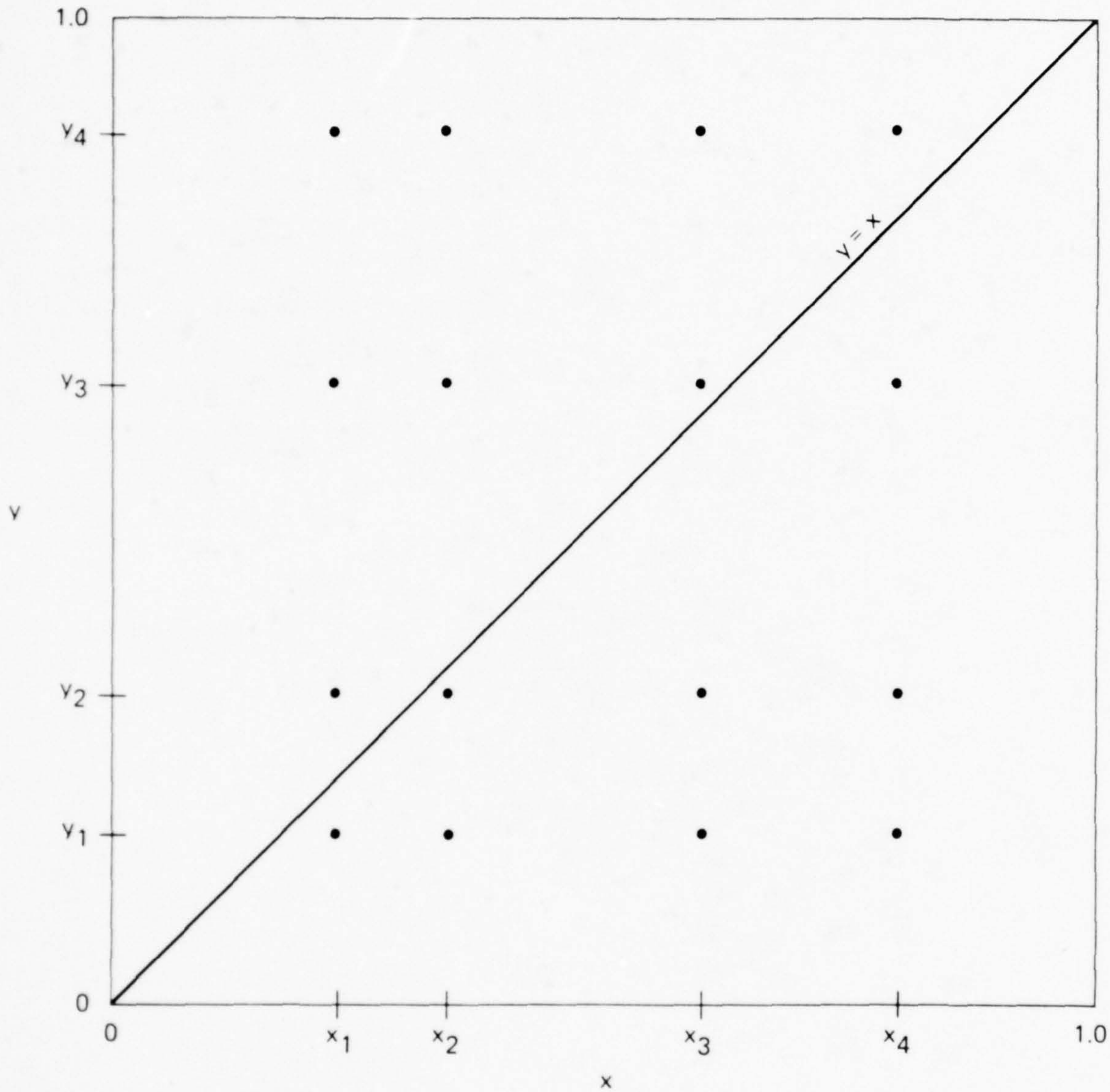


Figure 4-2

LATTICE POINTS IN SINGLE FACTORIAL EXPERIMENTAL DESIGN

them is consistent with the representation. It is easy to infer the order of the rows and columns from the indifference curves. One merely has to project the indifference curves through the points in the lattice onto the line, $\underline{y} = \underline{x}$. The ordering of the rows and columns is the same as the ordering of the indifference curves on that line. This example illustrates the conditions under which a rank ordering is Min-representable. Simply stated, the lowest ranked objects must be in the same row or column; when that column or row is removed, the lowest ranked objects in what remains should also be in the same row or column, and so on. A more formal statement of this property is stated in the following theorem.

<u>Theorem</u>. Let $\underline{A}_1 = \{\underline{a}_{11}, \underline{a}_{12}, \ldots, \underline{a}_{1m}\}$ and $\underline{A}_2 = \{\underline{a}_{21}, \underline{a}_{22}, \ldots, \underline{a}_{2n}\}$, and let $\succsim$ be a weak ordering on $\underline{X} = \underline{A}_1 \times \underline{A}_2$. $\succsim$ is Min-representable if any of the following conditions holds:

1. $\underline{m} = \underline{n} = 1$.
2. There exists $\underline{\ell} \varepsilon \{1, 2, \ldots, \underline{m}\}$ such that $\underline{x}_{ij} \succsim \underline{x}_{\ell k}$ for all $1 \leq \underline{i} \leq \underline{m}$, $\underline{i} \neq \underline{\ell}$, for all $1 \leq \underline{j},\underline{k} \leq \underline{n}$, and
   $\succsim$ restricted to $(\underline{A}_1 - \{\underline{a}_{1\ell}\}) \times \underline{A}_2$ is Min-representable.
3. There exists $\underline{\ell} \varepsilon \{1, 2, \ldots, \underline{n}\}$ such that $\underline{x}_{ij} \succsim \underline{x}_{k\ell}$, for all $1 \leq \underline{j} \leq \underline{n}$, $\underline{j} \neq \underline{\ell}$, for all $1 \leq \underline{i},\underline{k} \leq \underline{m}$, and
   $\succsim$ restricted to $\underline{A}_1 \times (\underline{A}_2 - \{\underline{a}_{2\ell}\})$ is Min-representable.

A similar theorem for Max-representable orders may be obtained by reversing the order of the empirical inequality, $\succsim$, in conditions 2a and 3a.

Before an outline of a proof is presented, two points should be made about the theorem. First, the conditions presented are sufficient for Min-representability, but not necessary. If the $\underline{k}_i$ and $\underline{\ell}_j$ assigned to the rows and columns are assumed to be unique, then the conditions are both necessary and sufficient. The problem occurs when a row and column have the same value. In this case, the row and column are on the same indifference curve, and the ordering among the objects in the row and column is random. When this occurs, the row and column must be treated as a unit, a procedure which is not considered in the theorem. A second point is that the theorem rests on the fact that both $\underline{A}_1$ and $\underline{A}_2$ are finite sets. A substantially different theorem and proof would result if one of these sets were infinite.

A formal proof of this theorem will not be given here. Rather, the procedure is outlined by which a representation may be found for a rank ordering on an $\underline{m}$ x $\underline{n}$ matrix using the conditions of the theorem. The procedure contains the following steps:

Step 1: Find the row or column containing the lowest ranks. Conditions 2a and 3a assure us that such a row or column exists. (We are assuming that either $\underline{m}$ or $\underline{n}$ is greater than 1.)

Step 2: Assign the stimuli representing that row or column the value $\frac{1}{\underline{m} + \underline{n}}$, and eliminate it from the matrix.

Step 3: Repeat Step 1. This time, assign the lowest row or column the value $\frac{2}{\underline{m} + \underline{n}}$, and eliminate it from the matrix.

Step 4: Continue the above procedure, and assign rows and columns values in increments of $\frac{1}{\underline{m} + \underline{n}}$, until a single cell remains in the matrix. (Conditions 2 and 3 assure that this may be done). At this point, all rows and columns except one row and one column have been assigned values consistent with the Min function. The remaining row and column may be assigned the values $\frac{\underline{m} + \underline{n} - 1}{\underline{m} + \underline{n}}$ and 1, in arbitrary order.

The ordering of the minima of the assigned row and column values is consistent with the empirical ordering when consistency is defined as above. An example of the application of this procedure to a 4 x 4 matrix is given in Figure 4-3.

In this example, the bottom row contains the lowest ranks, and that row is consequently assigned the value 1/8. Figure 4-3b shows the matrix with the scale values replacing the ranks in the bottom row. Of the part of the matrix remaining, the first column contains the lowest ranks. This column is assigned the value 2/8 and the matrix with scale values replacing the ranks in the first column is shown in Figure 4-3c. The above process continues until a single cell remains in the matrix. The row and column of this cell are arbitrarily assigned the values 7/8 and 1. The completely scaled matrix is shown in Figure 4-3d.

The procedure for finding a Max-representation for a rank ordering is completely analogous to the above procedure, with the exception that rows and columns are eliminated from the highest-ranked to the lowest-ranked instead of the reverse order.

12	5	4	1
11	6	3	2
10	9	8	7
15	14	16	13

a. Original rank order

	12	5	4	1
	11	6	3	2
	10	9	8	7
1/8	1/8	1/8	1/8	1/8

b. Matrix after first assignment of value

	2/8			
	2/8	5	4	1
	2/8	6	3	2
	2/8	9	8	7
1/8	1/8	1/8	1/8	1/8

c. Matrix after second value assignment

	2/8	4/8	5/8	7/8
8/8	2/8	4/8	5/8	7/8
6/8	2/8	4/8	5/8	6/8
3/8	2/8	3/8	3/8	3/8
1/8	1/8	1/8	1/8	1/8

d. Completely scaled matrix

Figure 4-3

ILLUSTRATION OF SCALING PROCEDURE FOR MIN REPRESENTATION

4.1.2 <u>Consistency of Min- and Max-representations</u> - In order for the functions Min and Max to represent a person's judgments of relative truth, three conditions must be met:

1. The judgments of conjunctions must be Min-representable

2. The judgments of disjunctions must be Max-representable

3. The Min- and Max-representations must be compatible.

The conditions by which Min- and Max-representability were assessed were given in the last section. This section describes the conditions necessary for compatibility of the two orders.

The Min and Max scaling procedure yields a partial order on the rows and columns of the matrix. This order is not a complete order, because some of the choices are arbitrary. The procedure given above shows that arbitrary assignments were made for the last row and column of the matrix. In fact, there is somewhat less certainty about the inferred rank ordering. This ordering is arbitrary for all rows and columns assigned values after the point when the portion of the matrix remaining consists of a single row or column. Since the Min scaling procedure starts with the lowest ranks, the highest-ranked rows and columns will not be completely ordered. Similarly, the lowest-ranked rows and columns are not completely ordered by the Max scaling procedure. These rank orders may be compared to the extent that they are complete. A measure of the extent of incompatibility between the two orderings is the number of reversals in the partial ordering of the Max-representation necessary to make it consistent with the Min-representation.

Figure 4-4 illustrates the manner in which compatibility may be investigated. In the figure, two rank-orderings are presented on the same matrix, with their respective Min- and Max-representations, and the partial ordering of rows and columns. In this case, it is apparent that the Min- and Max-representations are compatible.

4.1.3 <u>Ordinal conditions for product and inverse product representations</u> - Both the product relationship for conjunction and the inverse product relationship for disjunction imply an overall multiplicative relationship which must be satisfied by the rank ordering. In the case of conjunction, this relationship is straightforward. In the case of disjunction, the order is inversely related to the product of the complements of the truth values. Since, under both cases, the obtained values are non-negative, the ordering must satisfy the conditions of additive conjoint measurement in order for it to be representable by either a product rule or an inverse product rule. Necessary and sufficient conditions for an additive representation are presented elsewhere (Krantz, Luce, Suppes, & Tversky, 1971), and will be briefly mentioned here.

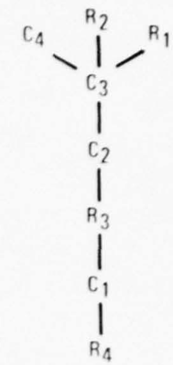
The two basic empirical conditions a rank ordering must satisfy for representation by an additive function are independence and cancellation. Independence states that the ordering of the cells in a row of the matrix is the same for all rows; similarly, the ordering within each column is the same for all columns. This condition is very important because it is necessary for any representation in which the value of the combination of objects is a strictly monotonic function of the value of the individual objects. (Of course, Min and Max are not strictly monotonic functions.) Failure of an ordering to satisfy independence rules out a great number of functions in addition to additive functions. Cancellation

12	5	4	1	R_1
11	6	3	2	R_2
10	9	8	7	R_3
15	14	16	13	R_4
C_1	C_2	C_3	C_4	

a. Min-representable rank order

	2/8	4/8	5/8	7/8	
8/8	2/8	4/8	5/8	7/8	R_1
6/8	2/8	4/8	5/8	6/8	R_2
3/8	2/8	3/8	3/8	3/8	R_3
1/8	1/8	1/8	1/8	1/8	R_4
	C_1	C_2	C_3	C_4	

b. Obtained representation of ordering



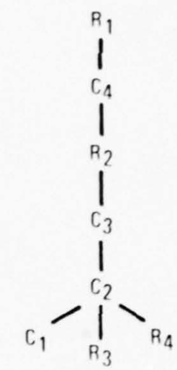
c. Partial ordering of rows and columns inferred from rank order.

3	1	2	4	R_1
10	8	9	5	R_2
15	13	12	6	R_3
16	14	11	7	R_4
C_1	C_2	C_3	C_4	

d. Max-representable rank order

	1/8	4/8	5/8	7/8	
8/8	8/8	8/8	8/8	8/8	R_1
6/8	6/8	6/8	6/8	7/8	R_2
3/8	3/8	4/8	5/8	7/8	R_3
2/8	2/8	4/8	5/8	7/8	R_4
	C_1	C_2	C_3	C_4	

e. Obtained representation of ordering



f. Partial ordering of rows and columns inferred from rank order

Figure 4-4

ILLUSTRATION OF COMPATIBILITY OF MIN AND MAX REPRESENTATIONS

is a set of conditions assuring that the ordering obtained is consistent with an additive rule in particular. The number of cancellation conditions which must be satisfied depends on the number of stimuli being ranked. A rank order satisfying independence and cancellation may be represented by an additive function of the rows and columns. Since the truth values are assumed to be positive, the obtained additive scale values may be taken to be the logarithms of the underlying scale values.

4.1.4 Consistency of product and inverse product representations - In order for orderings on conjunctions and disjunctions to be simultaneously consistent with product and inverse product rules, respectively, it must be possible to generate one of the orders using the scale values obtained from the other ordering. That is, it must be possible to take the scale values for the conjunctive ordering, subtract them from 1, find their products, and obtain the reverse of the ordering of disjunctions. Thus, the scale values, in addition to satisfying the linear inequalities implied by the individual rank orderings, must satisfy a set of non-linear equations implied by consistency of the two representations. Necessary and sufficient empirical conditions for two rank orders to simultaneously satisfy the inequalities and equations stated above are of necessity quite complicated. Consequently, consistency was tested using a general non-linear optimizing computer program.

The methods of this experiment may be compared with those of Oden (1977b) and Thole et al. (1979). Three major differences appear between these methods and those of the previous studies. First, this experiment involves a test based on an ordinal relationship which has already been validated. Second, this experiment involves the analysis of the data of individual subjects. Third, this experiment

simultaneously tests the agreement of the ordering of conjunctions and the ordering of disjunctions with a particular representation.

4.2 Method

4.2.1 Subjects - Subjects were 11 midshipmen from the U.S. Naval Academy, Annapolis, Maryland, and 3 employees of DDI, McLean, Virginia. The DDI employees had participated in Experiment 1; the midshipmen were totally unfamiliar with this line of research. The experiment lasted approximately 1 hour. The midshipmen participated without pay; the DDI employees received no compensation other than their regular salary.

4.2.2 Stimuli - The two sets of sentences labeled Set A and Set B in Table 4-1 were used to construct the stimulus sentences for this experiment. Compound sentences evaluated by subjects were formed by the factorial combination of these sets using the connectives "and" and "or". The simple sentences were chosen from those used in Experiment 1. Several criteria were used to decide which sentences were used in this experiment. The direct ratings of the sentences in Experiment 1 were used to guide the choice of sentences for this experiment. Sentences were chosen so that they covered a wide range from quite true to quite false. The sentences which appeared truest and least true were placed in the same set. (This was done to avoid the occurrence of lexicographic orderings in the subjects' responses, because lexicographic orderings do not distinguish between the two models.) Within each set, the sentences were to represent a roughly even distribution of truth values.

A total of 16 conjunctions and 16 disjunctions were formed by using all combinations of a single sentence from each set. For the first part of the experiment, each

Sentences Used to Construct
Experimental Stimuli

Set A

Los Angeles has clean air.

A man five feet, eleven inches tall is very tall.

Medicine tastes terrible.

It is uncomfortable to wear your left shoe on your right foot.

Set B

A lemon is a large fruit.

Four inches is a heavy snowfall.

Chevrolet makes excellent cars.

Jockeys are very short.

Table 4-1

compound sentence was typed on a separate card; the subjects rearranged the cards to represent the rank order. For the direct estimates, all 16 sentences of a given type were placed in a random order on a single sheet of paper.

4.2.3 Procedure - Subjects were initially given a packet containing all instructions, cards, and forms. The first page of the packet contained the following general instructions:

> Often in real life, we encounter statements which seem neither entirely true nor entirely false. Given two of these statements, we may be able to say that one or the other is more true, although we could not say that either is absolutely true or absolutely false.
>
> In the three parts of this experiment, you will be making two kinds of estimates about the truth of sentences. In the first two parts, you will be rank-ordering several sentences from the most true to the least true. In the third part of the experiment, you will be assigning a numerical estimate of the truth of the sentences.
>
> There are no right or wrong answers in this experiment, and this test in no way measures your intelligence or personality. In fact, your responses will not be compared directly to the responses of any other subject.
>
> Specific instructions are given with each part of the experiment. Read these instructions carefully before you work on that part. If you do not understand the instructions, ask one of the experimenters before you begin. Work quickly, but carefully. Your help in this research is much appreciated.

After subjects had finished reading the general instructions, they began the experiment. All subjects completed the parts of the experiment in the same order. The first part of the experiment was the ranking of the conjunctions, followed by the ranking of the disjunctions. This was followed by the direct estimates of the conjunctions, disjunctions, and simple sentences, in that order. The order was fixed for several reasons: the rank-order tasks were

placed before the direct estimate tasks because assigning numbers to sentences might interfere with the ordering task; conjunctions were completed before disjunctions because the evaluation of a conjunction was a considerably easier task than the disjunctions; and the simple direct estimates were placed last so that subjects would not feel constrained to use these ratings in some simple way to obtain the ratings of the compound sentences.

Subjects received the following instructions for the first part of the experiment, which was the ranking of conjunctions.

> Attached to these instructions is a packet containing sixteen cards. A compound sentence is written on each of the cards. Your task for this part of the experiment is to read the sentences and arrange the cards so that the truest sentence is on the top card, the next truest sentence is underneath it, and so on until you get to the least true sentence, which will be on the bottom card. When you are finished, the cards will be arranged in a pile from the truest on top down to the least true. Attach the paper clip to the pile of cards and proceed to Part II.
>
> It is easy to get confused on this task, so work carefully. When you are done, go through the stack of cards from top to bottom, checking to make sure that each sentence is truer than those on cards closer to the bottom of the stack and not as true as those on the cards closer to the top of the stack.

The second part of the experiment involved the rank ordering of the disjunctions. Subjects were told that their task was the same as in Part I, although the sentences they would be evaluating were slightly different. When subjects had completed the second part, they started on the direct estimates. In this part of the experiment, subjects read each sentence, and assigned to it a number between 0 and 100 representing its perceived truth. Thus, a completely true sentence was to receive the score 100, a

completely false sentence the score 0, and a sentence of intermediate truth a number in between in proportion to its perceived truth. Subjects rated the 16 conjunctions, followed by the 16 disjunctions, followed by the 8 simple sentences. When all sentences had been rated, the experiment was concluded. Data consisted entirely of the rankings and the direct estimates.

4.3 Results

4.3.1 Consistency - For the two types of compound sentence, the consistency of the subjects was measured by the product-moment correlation between the direct estimate of a sentence pair and the number of sentences over which the item was judged more true in the rank order. These correlations are presented in Table 4-2. The median correlation was .86 for the conjunctions and .83 for the disjunctions. Subject 4 did not complete the direct estimates of the disjunctions; direct estimate data from this subject are included only in the analyses of conjunctions. The sum of the correlations for the conjunctions and disjunctions was taken to be a measure of overall consistency. The overall consistency measures are presented in Table 4-2. The median level of consistency was 1.68.

4.3.2 Analysis of ordinal data - The extent to which a rank order may be represented by a specific combination rule may be measured by the extent to which it must be changed in order to satisfy the empirical conditions implied by the rule. For the product and inverse product rule, the conditions are those given by the axioms for additive conjoint measurement. Thus, the extent to which an order deviates from the product rule may be measured by the minimum number of adjacent pairwise reversals in the rank order necessary for the ordering to be additive. For the Min and Max rules,

Consistency of Rank Order and Direct Estimates

Subject	Correlation		Sum
	Conjunctions	Disjunctions	
1	.75	.72	1.47
2	.90	.95	1.85
3	.91	.82	1.73
4	.45	--	--
5	.77	.53	1.30
6	.85	.96	1.81
7	.71	.36	1.07
8	.88	.85	1.73
9	.86	.81	1.67
10	.94	.83	1.77
11	.86	.75	1.61
12	.79	.86	1.65
13	.86	.83	1.69
14	.93	.94	1.87

Table 4-2

the measure is the number of pairwise reversals to the nearest Min-representable or Max-representable ordering. The numbers of reversals for the two sets of combination rules are given in Table 4-3. Subtotals are listed for both the more consistent and the less consistent half of the subjects.

As can be seen from the table, the additive models are somewhat closer to the actual rank orders than the Min or Max rules. In $\underline{t}$-tests there were significantly fewer deviations from the multiplicative (product and inverse product) models than from the extreme value (Min and Max) models for the disjunctive ordering, $\underline{t}(13) = 2.36$, $\underline{p} < .05$, and overall for both orderings, $\underline{t}(13) = 2.47$, $\underline{p} < .05$. The difference is in the same direction, but not significant for the conjunctive orderings, $\underline{t}(13) = 1.74$. For the most consistent half of the subjects, the difference was significant for both the conjunctive and disjunctive orderings, $\underline{t}(13) = 2.73$ and 3.64, respectively, $\underline{p} < .05$, as well as overall, $\underline{t}(13) = 3.88$, $\underline{p} < .01$. Table 4-3 indicates that for the more consistent half of the subjects, no rank order was better fit by the extreme value than by the multiplicative model. The probability of such an extreme result calculated from the binomial distribution with $\underline{p} = 0.5$ is .06, .03, and .02 for the conjunctive, disjunctive, and combined orderings, respectively. For the less consistent subjects, however, neither model fit particularly well. The two results--that obtained rank orderings are more closely fit by the appropriate additive model, and that higher consistency is associated with better fit of the multiplicative functions--give strong support for the product and inverse product combination rules.

The compatibility of the representations for the two orderings was checked for those subjects for whom both orderings completely satisfied one of the models. Specifically,

Pairwise Reversals in Rank Order Necessary to Satisfy Representation

Subject	Conjunctions		Disjunctions		Total	
	Product	Minimum	Inverse Product	Maximum	Multi-plicative	Extreme Value
14	0	0	0	9	0	9
2	2	10	4	11	6	21
6	0	9	0	10	0	19
10	0	0	1	3	1	3
3	6	8	0	5	6	13
8	4	12	0	0	4	12
13	3	5	3	4	6	9
Average	2.14	6.29	1.14	6.00	3.28	12.29
9	0	0	0	0	0	0
12	14	15	18	15	32	30
11	0	0	0	0	0	0
1	3	3	9	11	12	14
5	4	4	8	5	12	9
7	5	1	12	11	17	12
4	11	10	6	6	17	26
Average	5.29	4.71	7.57	8.29	12.86	13.00
Overall Average	3.71	5.50	4.36	7.14	8.07	12.64

Table 4-3

four subjects completely satisfied the multiplicative models and two subjects satisfied the extreme value models. For all four of the subjects who satisfied the multiplicative models, it was possible to find a set of truth values which simultaneously reproduced both the conjunctive and disjunctive orderings when combined with the product and inverse product rules. One of the two subjects (Subject 11) satisfying the extreme value models produced compatible orderings. For the other (Subject 9), there were eight inconsistencies between the partial orders inferred from the two rank orderings. These data considerably strengthen the evidence for the product and inverse product combination rules.

4.3.3 Analysis of direct estimates - In addition to estimating the truth of sentences which were conjunctions and disjunctions of simple sentences, subjects also rated the simple sentences directly. Thus, it is possible to assess how well the ratings of the compound sentences may be predicted from the simple ratings by using the combination rules in question. There are two parts to this assessment. The first issue is which rule is better at predicting the direct estimates; the second is whether either rule does a good job at prediction. For each appropriate rule the mean squared deviation of the direct estimates from the values predicted by applying the rule to the simple estimates was calculated. The two rules may be compared by an examination of these deviations. The second question was answered by comparing deviations from each of the combination rules to the deviations from the grand mean of the direct estimates. In order for a rule to be an adequate predictor of the direct estimates, it was judged that it should be able to predict those values better than a single number.

The mean squared deviations from the mean and the appropriate extreme value and multiplicative rules are

presented in Table 4-4. An analysis of this table shows no clear pattern in the values. For the conjunctions, there were no significant differences between Min and product rules in predicting the direct estimates. The mean was a significantly better predictor than either combination rule for a single subject. For the disjunctions, Max was significantly better than the inverse product for two subjects; for one subject this relationship was reversed. For three subjects, one of the combination rules predicted the direct estimates better than the mean did; for one subject this relationship was reversed.

The only reasonable conclusion from the data in Table 4-4 seems to be that neither model does particularly well at predicting the perceived truth of the compound sentences. This lack of fit may be due to inadequacies of the models themselves, or to response biases; i.e., subjects may use different processes to evaluate the compound sentences than they do on the simple sentences. Such response biases might lead simple sentences to be judged relatively more true than equivalent conjunctions or disjunctions. Although there are considerable discrepancies between the predictions of the models, the predicted values were highly correlated with the actual direct estimates. The median correlation between the predictions of the Min and product rules with the estimates of the conjunctions were .80 and .82, respectively. The median correlation between the Max and inverse product rules and the estimates of the disjunctions were .79 and .82, respectively. This result indicates that the lack of fit between the models and the data may be due to an incompatibility of the responses, rather than an incompatibility between the proposed and subjective combination rules. To avoid confounding these interpretations, estimates of the goodness of fit were made from the ratings of the compound sentences alone.

Mean Squared Deviations From Model Predictions
Taken from Simple Estimates

Subject	CONJUNCTIONS			DISJUNCTIONS		
	Mean df=15	Minimum df=16	Product df=16	Mean df=15	Maximum df=16	Inverse Product df=16
14	965.2	1361.0	1512.0	966.3	911.0	866.4
2	434.9	625.0	820.3	582.7	615.2	810.6
6	437.4	407.8	710.1	349.1	339.1	534.5
10	792.2	508.1	710.7	1008.0	661.6	698.4
3	1143.0	976.6	976.6	1073.0	1406.0	1406.0
8	743.3	1120.0	1218.0	876.7	382.8	480.5
13	826.7	731.2	813.8	905.0^{a}	0.0^{b}	37.5^{c}
9	266.7^{a}	1200.0^{b}	1565.0^{b}	281.6	317.2	411.1
12	600.0	487.5	670.9	936.2^{ab}	1294.0^{a}	469.6^{b}
11	266.7	243.8	460.5	500.0^{a}	131.2^{b}	370.5^{a}
1	641.7	610.9	729.7	726.2^{a}	329.7^{b}	415.6^{ab}
5	641.7	791.2	951.9	673.2^{a}	1452.0^{ab}	2005.0^{b}
7	478.2	779.7	779.2	706.7	825.0	1119.0
4	874.9	1336.0	1368.0	---	---	---

NOTE: Numbers in the same row with different superscripts are significantly different by F-test, $p < .05$.

Table 4-4

Least squares estimates of the row and column values were made for the rules in the following manner. For the extreme value models, rows and columns were eliminated in order of their average rating in a manner similar to the ordinal scaling method. The mean of the row was assigned to the row, which was then eliminated from the matrix. Estimates for the product and inverse product models were made using Tukey's one-degree-of-freedom test for nonadditivity. This test fits the equivalent of a multiplicative model. Both of these methods assume that the direct estimates are an interval-scale measure of perceived truth and may be transformed by any linear function. Consequently, both rules must be better predictors than the mean. Standard analysis of variance techniques may be used to test the significance of these improvements. Similarly, the ratio of the residuals from the appropriate two rules may be used to compare the accuracy of the two models.

The summaries of the analysis of variance for the conjunctions and disjunctions are presented in Tables 4-5 and 4-6, respectively. Table 4-5 shows that either combination rule provided a significant improvement in predictive power over the mean for most subjects. Specifically, the minimum rule predicted the estimates better for 9 subjects and the product rule was better than the mean for all subjects. For 10 of the 14 subjects, the product rule was a significantly better predictor than the minimum rule.

The results for the disjunctions are similar, although somewhat more complicated. Again, both models generally predicted the direct estimates significantly better than the mean. However, the comparison between the models was not as clear-cut as in the case of the conjunctions. For 5 subjects, the inverse product fit the data significantly better than the maximum. However, one subject (Subject 13) seemed definitely to be using the maximum rule as this

Summary of Analyses of Variance for

Direct Estimates of Conjunctions

Subject	Total SS df=16	Mean df=1	MINIMUM SS df=7	MINIMUM Residual df=8	PRODUCT SS df=7	PRODUCT Residual df=8	F df=(8,8)
14	67,033	52,556	11,598[h]	1,879.4	11,282[t]	194.8	9.65[h]
2	60,000	53,477	5,223[f]	1,300.3	6,263[t]	260.0	5.00[f]
6	62,375	55,814	4,417	2,143.8	5,602[h]	959.3	2.23
10	54,114	42,230	11,464[t]	419.7	11,788[t]	96.13	4.37[f]
3	45,625	28,477	14,804[h]	2,343.8	16,897[t]	250.8	9.35[h]
8	43,550	32,400	8,700[f]	2,450.0	10,819[t]	331.4	7.39[h]
13	56,500	44,100	10,002[f]	2,397.9	12,117[t]	222.9	10.76[h]
9	61,600	57,600	3,000	1,000.0	4,000[t]	0.0	∞[t]
12	34,600	25,600	7,408[f]	1,591.7	7,220[f]	1,780.0	0.894
11	82,400	78,400	3,000	1,000.0	4,000[t]	0.0	∞[t]
1	33,650	24,025	7,908[f]	1,716.7	9,438[t]	187.1	9.18[h]
5	51,959	42,333	6,601	3,024.8	8,317[h]	1,309.0	2.31
7	65,375	58,202	5,040	2,133.3	6,041[f]	1,132.0	1.88
4	62,074	48,952	10,304[f]	2,818.8	12,381[t]	742.1	3.80[f]

[f] $p < .05$

[h] $p < .01$

[t] $p < .001$

Table 4-5

Summary of Analyses of Variance for

Direct Estimates of Disjunctions

Subject	Total SS df=16	Mean df=1	MINIMUM SS df=7	MINIMUM Residual df=8	PRODUCT SS df=7	PRODUCT Residual df=8	F df=(8,8)
14	77,496	63,001	13,328[t]	1,166.5	14,463[t]	31.6	36.9[t]
2	57,969	49,229	6,917[f]	1,822.9	8,654[t]	86.2	21.1[t]
6	69,625	64,389	3,644	1,591.7	5,042[t]	193.5	8.23[h]
10	76,620	61,504	13,183[h]	1,932.7	13,698[h]	1,418.4	1.36
3	42,500	26,406	12,761[f]	3,333.3	14,689[h]	1,404.5	2.37
8	49,250	36,100	11,048[f]	2,102.1	12,295[t]	854.9	2.46
13	106,600	93,025	13,575[t]	0.0	13.109[t]	466.2	0.0[t]
9	61,225	57,002	3,320[f]	902.5	4,071[t]	151.5	5.96[f]
12	37,300	23,256	8,036	6,008.3	11,586[f]	2,458.1	2.44
11	103,600	96,100	7,200[t]	300.0	7,475[t]	24.5	12.2[t]
1	47,950	37,056	9,188[f]	1,706.3	9,516[h]	1,377.9	1.24
5	32,225	22,127	6,742	3,356.3	6,392	3,706.4	0.906
7	78,200	67,600	5,498	5,102.1	5,374	5,225.9	0.976

[f] $p < .05$

[h] $p < .01$

[t] $p < .001$

Table 4-6

rule exactly predicted the data for that subject. Furthermore, the scale values for this subject were the same as the estimates given by the subject for the simple sentences. Clearly, there is an indication of the possibility of individual differences in these data.

In summary, the results of the analysis of the direct estimates confirm those of the ordinal data in pointing to the multiplicative model as a descriptive model of judged truth of disjunctions and conjunctions of simple sentences. However, one of the subjects seemed to be using the maximum rule for the disjunctions, indicating the possibility of individual differences in this area. Also, there seems to be some difference in the processes used to assign a truth value to simple and more complex sentences.

4.4 Discussion

The general conclusion of this experiment is that the product and inverse product rules provide a reasonable account of the way in which people perceive the truth of conjunctions and disjunctions of simple sentences. The evidence for these combination rules comes from both the rank orders and direct estimates of the truth of conjunctions and disjunctions. Specifically, rank orders more closely fit the conditions of the multiplicative functions than the extreme value functions. One illustration of this difference is the fact that no subject's rank orders completely satisfied the extreme value model without also satisfying the multiplicative model. The differences between the models were accentuated for the more consistent subjects. Furthermore, the orderings for conjunctions and disjunctions were compatible with the same set of scale values for the multiplicative models, while this was not true for the extreme value models. The direct estimate data

were also better fit by the multiplicative models than by the extreme value models. However, the scale values which could be inferred from the direct estimates of the compound sentences were not the same as the direct estimates of the simple sentences. This result may indicate a problem with the multiplicative model, or it may indicate a response bias caused by the variation in semantic complexity between the simple and compound sentences. The evidence seems to favor the latter interpretation, but is quite weak on this issue. In addition, the direct estimates gave some indication of the possibility of individual differences in the function representing disjunctions.

On the surface, these results seem to correspond to those of Oden (1977b) and directly to contradict those of Thole et al. (1979). However, a closer examination of the results indicates that the results of this experiment are consistent with both of these previous studies. These results also shed light on how these seemingly contradictory findings may be reconciled. In comparing the results of these experiments, it is important to keep in mind the methodological differences between them. Chief among these differences is the fact that this experiment tests the predictions of the two sets of combination rules on individual subjects using a previously validated rank order.

Stimulus material consisted of explicit conjunctions and disjunctions of sentences in this experiment, as well as those of Oden. However, the sentences in this experiment cover a much wider range of topics, rather than the statements of set inclusion used by Oden. The inclusion of such a variety of sentences may have increased the difficulty of the task for the subject; however the data gave evidence for the same combination rules. In fact, the results of this experiment are stronger than those of Oden for two reasons:

1) The analysis was based on ordinal relations among the sentences; and 2) this analysis tested whether the orderings on the conjunctions and disjunctions simultaneously satisfied the extreme value rules or the multiplicative rules. On the second point, all subjects whose rank orderings satisfied the product and inverse product models individually, satisfied them simultaneously.

The results of Thole et al. (1979) are also consistent with those reported here. Specifically, their finding that neither the minimum nor product combination rule predicts the membership value of the intersection of two sets, when applied to the values of the individual sets, corresponds to our findings summarized in Table 4-4. The agreement highlights the response biases which may occur when rating sentences which differ in complexity. Interestingly, the finding of Thole et al. that the minimum was slightly closer to the actual values than the product also agrees with the corresponding analysis in the current data. The total deviations (taken from Table 4-4) from the minimum rule were less than those from the product (11,179 versus 13,287). However, a more fine-grained analysis of the direct ratings of the compound statements themselves shows that this small difference does not reflect the processes underlying the data. Thus, the response bias masked the compensatory processes which could be identified by further analysis.

The tested conditions have a strong impact on the nature of the representation of partial truth. In particular, it should be possible to use combinations of statements to derive a scale of truth which has properties stronger than ordinal. The exact uniqueness of this scale is quite complex. The individual orderings of the conjunctions and disjunctions force the logarithm of the scale to be an ordered metric scale. The requirement for compatibility of the two scales imposes a much more complicated constraint on

the uniqueness of the scales. One area in which further work would be useful is the development of representation and uniqueness theorems for reflecting the structure of partial truth judgments.

It is important to note that while the obtained results favor a multiplicative rule as a descriptive model of individual judgments, extreme value rules might be more justified in a normative setting. For example, the multiplicative rule implies that the truth of the conjunction of a statement with itself is less than the truth of the statement. This relationship may be considered an unsatisfactory guide for behavior, and may indicate that the rules which appear to be the natural rules in human judgment should not be used in a normative setting. In this case, it might be difficult for an analyst to assess individual truth or measurement functions without first developing methods to reconcile this discrepancy. Further research assessing the difficulties and benefits of using fuzzy set theory in decision analysis would serve both to increase the availability and usefulness of fuzzy set theory as a decision aid, and to enlarge our understanding of the process by which individuals understand imprecise information.

Several areas which would benefit from future research have been identified. The first involves further theoretical and empirical research on the axiomatic foundations of imprecise reasoning, to develop a firmer understanding of the measurement and scaling implications of the approximate procedures individuals use to understand complex systems. The second area involves further empirical investigations on imprecise reasoning in more meaningful and realistic settings, in order to increase the ecological validity of research in the area, and to improve understanding of the ways in which semantic constraints interact with the rules studied here.

Finally, the effectiveness of fuzzy set theory as a decision aid should be investigated; this area probably has the greatest potential, as it would contribute to improving both our descriptive understanding of human reasoning processes, and our ability to assist individuals in making effective decisions.

SUPPLEMENTAL DISTRIBUTION LIST

5.0 DISCUSSION AND CONCLUSIONS

The major goal of this research program has been to investigate the correspondence between the rules of fuzzy set theory and the structure of human judgments. In this context, the results have been quite favorable for fuzzy set theory. The results of the first experiment indicate that individuals can process imprecise information in a consistent and transitive manner. Because of individual consistency, it is possible to construct an ordinal measure of partial truth. The second experiment investigated the consistency of judgments of relative truth for more complex sentences--in particular, the rules by which the partial truth values of simple sentences were integrated when these sentences were combined using the logical operations of conjunction and disjunction. Two candidate sets of combination rules were examined: multiplicative rules and extreme value rules. These rules represent two alternative formulations of fuzzy set theory, each having strong implications about the nature of measurement scales of partial truth. The results of the experiment gave strong support for the multiplicative rules, and furthermore indicated the possibility of response biases in evaluating the truth of sentences with different degrees of complexity.

The results of these experiments influence both normative and descriptive theory. The most obvious impact, of course, is in the descriptive area, in the form of a firm empirical base for the use of fuzzy set concepts and a quantitative characterization of individual integration rules. The normative implications are more complex and consequently also more interesting: Although individual judgments of partial truth seem to be integrated by multiplicative rules, these rules may seem inappropriate from

a normative viewpoint. One major problem which may occur with the multiplicative rules involves the truth of the conjunction of a sentence with itself. If one requires that the truth of the conjunction of a sentence with itself is the same as the truth of that sentence, then the extreme value rules are the only permissible ones which satisfy the definitions for conjunction and disjunction. There is some question about whether the above stipulation should be a condition of rational behavior; nevertheless, there is a potential for problem in this area. Since some of the major areas of application of fuzzy set theory are in decision analysis (Watson, Weiss, & Donnell, 1979), policy capturing (Weiss, 1978), and other areas in which the theory is used to aid in decisionmaking and problem solving, great benefits may be obtained from future research in the area of applications of the theory.

It is our opinion that, in the near term, the most important research in this area will involve the evaluation of fuzzy set theory as a tool for aiding decisions. Fuzzy set theory has great potential in decision analysis in that it allows an analysis to proceed with imprecise inputs. Thus, the decisionmaker will be able to express his values and uncertainties in either vague or precise terms. However, assessment and use of the theory as a decision aid depend on establishing some correspondence between the normative laws of behavior involving imprecise concepts and the natural rules of human information processing. In order to evaluate the effectiveness of the theory as a decision aid, the following tasks must be accomplished:

- develop normative rules for the processing of imprecise information;
- use these rules to design decision aids; and

- compare the effort involved in using these fuzzy decision aids, the quality of the results, and the satisfaction of the decisionmaker with the whole process, with corresponding measures for more traditional methods of decision analysis.

The benefits of making decision analysis more responsive to the needs of the decisionmaker are substantial, and the room for improvement is great; hence, there is a potentially large payoff for future research in fuzzy set theory and its application to decisionmaking.

REFERENCES

Anderson, N. H. Information integration theory: a brief survey. In D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes (Eds.), Contemporary Development in Mathematical Psychology (Vol. 2). San Francisco: Freeman, 1974.

Dreyfuss, G. R., Kochen, M., Robinson, J., & Badre, A. N. On the psycholinguistic reality of fuzzy sets. In R. E. Grossman, L. J. San, & T. J. Vance (Eds.) Functionalism. Chicago: University of Chicago Press, 1975, pp. 135-149.

Goguen, J. A. The logic of inexact concepts. Synthese, 1969, 19, 325-373.

Hersh, H. M. & Caramazza, A. A fuzzy set approach to modifiers and vagueness in natural language. Journal of Experimental Psychology: General, 1976, 105, 254-276.

Kendall, M. G. Rank Correlation Methods. New York: Hafner Publishing Company, 1955.

Kochen, M. & Badre, A. N. On the precision of adjectives which denote fuzzy sets. Journal of Cybernetics, 1974, 4, 49-59.

Krantz, D. H.; Luce, R. D.; Suppes, P.; and Tversky, A. Foundations of Measurement (Vol. 1). Additive and polynomial representations. New York and London: Academic Press, Incorporated, 1971.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 1957, 54, 358-368.

Macvicar-Whelan, P. J. Fuzzy sets, the concept of height, and the hedge "very". IEEE Transactions on Systems, Man, and Cybernetics, 1978, SMC-8, 507-511.

Oden, G. C. Fuzziness in semantic memory: Choosing exemplars of subjective categories. Memory and Cognition, 1977, 5, 198-204. (a)

Oden, G. C. Integration of fuzzy logical information. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 565-575. (b)

Rosch, E. H. Cognitive representation of semantic categories. Journal of Experimental Psychology: General, 1975, 104, 192-233.

Thole, U.; Zimmermann, H. J.; & Zysno, P. On the suitability of minimum and product operators for the intersection of fuzzy sets. Fuzzy Sets and Systems, 1979, 2, 167-180.

Tukey, J. W. Exploratory Data Analysis. Reading, Massachusetts: Addison-Wesley Publishing Company, 1977.

Watson, S. R., Weiss, J. J., & Donnell, M. L. Fuzzy decision analysis. IEEE Transactions on Systems, Man, and Cybernetics, 1979, SMC-9, 1-9.

Weiss, J. J. Fuzzy Policy Capturing as a Decision Aid (Interim Report PR 78-30-83). McLean, VA: Decisions and Designs, Inc., 1978.

Zadeh, L. A. Fuzzy sets. Information and Control, 1965, 8, 338-353.

Zadeh, L. A. Probability measures of fuzzy events. Journal of Mathematical Analysis and Applications, 1968, 23, 421-427.

Reference Note 1

Hersh, H. M. Fuzzy reasoning: The integration of vague information. Unpublished manuscript, 1978. (Available from Harry M. Hersh, Digital Equipment Corp., 146 Main Street, ML3-2/E41, Maynard, MA 01754.)

APPENDIX A
LIST OF STIMULUS SENTENCES

LIST OF STIMULUS SENTENCES

1. Ohio is in the Midwest.
2. It's a long way from Pennsylvania to Canada.
3. Physicians make a lot of money.
4. A penguin is like a fish.
5. California is in the northern United States.
6. Four inches is a heavy snowfall.
7. A porpoise is a typical mammal.
8. Four pancakes is a lot for a lumberjack's breakfast.
9. A man five feet, eleven inches tall is very tall.
10. Mathematics is a difficult subject.
11. It is inappropriate to be barefoot in a restaurant.
12. Shoplifting is a serious crime.
13. Mexican food is very spicy.
14. Jockeys are very short.
15. President Carter has a strong accent.
16. Hamsters are popular pets.
17. It is uncomfortable to wear your right shoe on your left foot.
18. An umbrella is effective in keeping one dry in the rain.
19. Many stars are visible in the city sky at night.
20. It is not expensive to call someone in Europe from the United States.
21. Most Americans are overweight.
22. Cheesecake is very sweet.
23. Athletes are overpaid.

24. Alcohol excites people.

25. Ringo Starr (of the Beatles) is a great drummer.

26. It is unsafe to live in large American cities.

27. Bob Hope is very funny.

28. The Internal Revenue Service (IRS) is reasonably fair.

29. The United Nations is an effective organization.

30. Most lawyers act as though they have no conscience.

31. Los Angeles has clean air.

32. Chevrolet makes excellent cars.

33. White lies do no harm.

34. Medicine tastes terrible.

35. Women and men are treated equivalently in our society.

36. Greece is a nice place to go for a vacation.

37. Lake Erie is a large body of water.

38. The Olympics is an example of international cooperation.

39. Rhode Island has a lot of people.

40. West Germany is a highly industrialized country.

41. Fish have a strong odor.

42. The color of goldfish is yellow.

43. Skiing is a dangerous sport.

44. Tennis balls are fuzzy.

45. A lemon is a large fruit.

46. Smoking tobacco is beneficial to one's health.

47. Alaska is north of Canada.

48. A 35-year-old man is middle-aged.

49. Denver is near the Pacific Ocean.

50. Potato chips are salty.

APPENDIX B

A BRIEF EXPLANATION OF "STEM-AND-LEAF" DIAGRAMS

A BRIEF EXPLANATION OF "STEM-AND-LEAF" DIAGRAMS

The "stem-and-leaf" diagram is a recently developed means of displaying raw or calculated data in a manner which permits either rapid visual scanning or detailed study. It is essentially a kind of histogram, where equal intervals are labeled along the left-hand margin, and each occurrence of a value within a given interval corresponds to one token on the appropriate line of the main figure. Typically, the left-margin label will be either the interval's range or the lower bound of that range (choice of scale is at the pleasure of the user, but is typically designed to segment the sample into groups of no more than ten or fifteen). The tokens themselves are the least significant digits of the corresponding numbers. For example, to represent the set of observations {2, 4, 7, 7, 12, 15, 17, 19, 21, 28, 43}, we might use the five intervals 0-9, 10-19, 20-29, 30-39, and 40-49. Thus, the stem-and-leaf diagram would look like this:

0	2	4	7	7
10	2	5	7	9
20	1	8		
30				
40	3			

The reader can then read the diagram either as a gross histogram of the scores (four in the 0-9 interval, four in the 10-19 interval, etc.), or else look in more detail at the specific scores (for example, the entry "7" on the line labeled "10" corresponds to the observation "17" that was listed in the original sample.

As a further example, suppose you must represent given the scores: .0037, .0042, .0059, .0059, .0062, .0064, .0076, .0078, .0091, .0099, .0120, .0145, .0187, .0277, .0291. The stem-and-leaf representation of this set of data might look like this:

.0000-.0049	37	42						
.0050-.0099	59	59	62	64	76	78	91	99
.0100-.0149	20	45						
.0150-.0199	87							
.0200-.0249								
.0250-.0299	77	91						

Here again, the essential shape of the distribution may be read simply and directly, while each individual data item is also represented with no loss of information (for example, the entry "91" on the ".0050-.0099" line represents the observed value of .0091).

This simple but useful format is gaining popularity as a means of displaying data, and the reader should find it worth the small effort it takes to get comfortably acquainted with it.

APPENDIX C

DISTRIBUTION LIST

CONTRACT DISTRIBUTION LIST

Director Engineering Psychology Programs Code 455 Office of Naval Research 800 North Quincy Street Arlington, VA 22217	5 copies
DCASMA Baltimore Hampton Plaza Building, Room 200 300 East Joppa Road Towson, MD 21204	1 copy
Director Naval Research Laboratory Technical Information Division Code 2627 Washington, DC 20375	6 copies
Defense Documentation Center Cameron Station, Bldg. 5 Alexandria, VA 22314	12 copies
Commanding Officer ONR Branch Office Attn: Dr. J. Lester Building 114, Section D 666 Summer Street Boston, MA 02210	1 copy

SUPPLEMENTAL DISTRIBUTION LIST

CDR Paul R. Chatelier
Military Assistant for Training and Personnel Technology
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Director
Analysis and Support Division
Code 230
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Branch Office
Attn: Dr. C. Davis
536 South Clark Street
Chicago, IL 60605

Director
Statistics and Probability Program
Code 436
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Branch Office
Attn: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco 96503

Dr. Robert G. Smith
Office of the Chief of Naval Operations, OP987H
Personnel Logistics Plans
Washington, DC 20350

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-OOT
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Navy Personnel Research and Development Center
Manned Systems Design, Code 311
San Diego, CA 92152

Navy Personnel Research and Development Center
Code 305
San Diego, CA 92152

Navy Personnel Research and Development Center
Management Support Department
Code 210
San Diego, CA 92152

Mr. J. Barber
HQS, Department of the Army
DAPE-PBR
Washington, DC 20546

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director
Organizations and Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Edgar M. Johnson
Organizations and Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

U.S. Air Force of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Dr. Gordon Eckstrand
AFHRL/ASM
Wright-Patterson AFB, OH 45433

Dr. Stephen J. Andriole
Director
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Judith Daly
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Alphonse Chapanis
Department of Psychology
The John Hopkins University
Charles and 34th Streets
Baltimore, MD 21218

Dr. Meredith P. Crawford
Department of Engineering Administration
George Washington University
Suite 805
2101 L Street, NW
Washington, DC 20037

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Ward Edwards
Director, Social Science Research Institute
University of Southern California
Los Angeles, CA 90007

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Room 201
Boulder, CO 80309

Dr. Ronald Howard
Department of Engineering-Economic Systems
Stanford University
Stanford, CA 94305

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, NW
Washington, DC 20036 (3 copies)

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, CA 93017

Dr. Gary McClelland
Institute of Behavioral Sciences
University of Colorado
Boulder, CO 80309

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Menlo Park, CA 94025

Dr. Terence R. Mitchell
University of Washington
Seattle, WA 98195

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Professor Judea Pearl
Engineering Systems Department
University of California-Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

Professor Howard Raiffa
Graduate School of Business
Administration
Harvard University
Soldiers Field Road
Boston, MA 02163

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. David Dianich
Chairman, Department of
Business and Economics
Salisbury State College
Salisbury, MD 21801

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
UNITED KINGDOM

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsville, Toronto, Ontario
CANADA

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
UNITED KINGDOM

Dr. J. Baal Schem
Acting Director
Tel-Aviv University
Interdisciplinary Center for Technological
Analysis and Forecasting
Ramat-Aviv, Tel-Aviv
ISRAEL